FACULTY WORKING PAPER 91-0100

Universal Coalition-Proof Equilibrium

Bhaskar Chakravorti
Department of Economics

Charles M. Kahn
Department of Economics

Bureau of Economic and Business Research
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 91-0100

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

January 1991

Universal Coalition-Proof Equilibrium

Bhaskar Chakravorti

and

Charles M. Kahn

Department of Economics

We are grateful to George Mailath and Jacques Robert for helpful discussions on the subject. The authors retain responsibility for all errors.

UNIVERSAL COALITION-PROOF EQUILIBRIUM

ABSTRACT

We wish to characterize stable agreements in games in which coalitions can make non-binding self-enforcing deviations from such agreements. We allow for "universal" coalition formation, i.e. the validity of a deviation is checked not only against further deviations by subsets of the deviating coalition but also against deviations agreed upon by some members of the deviating coalition convincing players from the complementary coalition to deviate. The blocking device introduced here is a "threat" which is weaker than that of "trumping", associated with Bernheim, Peleg and Whinston's (1987) concept of Coalition-proof Equilibrium (CPE) and is stronger than an "objection", associated with an equilibrium concept (CCTE) based upon Greenberg's (1989) Coalitional Contingent Threats. We show that our solution concept has no logical relationship with either CPE or CCTE; however, the solution is a Nash Equilibrium refinement. We argue that the last result is quite unexpected since, a priori, a unilateral best-response property is not a necessary condition for stability (in a von Neumann-Morgenstern (1947) sense) in the presence of universal coalition formation. The key to resolving the information asymmetry inherent in our solution concept and its Nash refinement property is a "lateral induction" condition which is related to the idea of forwards induction in sequential games.

JEL Classification Number: 026

Address correspondence to:

Bhaskar Chakravorti
Department of Economics
University of Illinois at Urbana-Champaign
330 Commerce West
1206 South Sixth Street
Champaign, IL 61820

"Those who meet will meet again." —— *Zulu proverb.*

1. INTRODUCTION

Consider a finite n-player game in which coalitions of players communicate prior to actual play and make non-binding agreements on strategy choices. We wish to know which agreements are stable in such environments.

Following upon Aumann's (1959) "strong" Nash equilibrium, one answer to this question is given by Bernheim, *et al.*'s (1987) notion of Coalition-proof Nash equilibrium (CPE), a refinement of Nash equilibrium. This provides the starting point of our paper. To borrow Kalai's remark quoted in Greenberg (1989): "The concept of CPE does not go far enough in its analysis of stability. When considering a deviating coalition, the validity of the deviation is checked only against further deviations of subcoalitions of the deviating coalition. *(We refer to this as the "nestedness" assumption.)* However, members of the deviating coalitions could also deviate by convincing other players (from the nondeviating coalition) to deviate provided they improve their payoff *(we refer to this as universal coalition formation).*" The primary motivation of this paper is precisely to address Kalai's concerns and to characterize stable agreements in the presence of universal coalition formation. We show that the characterization is dramatically different from that of CPE.

This paper presents a new solution concept called Universal Coalition-proof equilibrium (UCPE). We propose the concept of UCPE as the natural non-cooperative counterpart of the consistent Bargaining Set (Dutta, *et al.* (1989)), a refinement of the Bargaining set (Aumann and

Maschler (1964), Mas-Colell (1989), Shapley and Shubik (1984)) in cooperative game theory. The UCPE contains the set of strong Nash equilibria but has no logical containment relationship with CPE. Moreover, there is no logical relationship with an equilibrium concept based on Greenberg's (1989, 1989a) notion of Coalitional Contingent Threats that also does not rely on the nestedness assumption. This latter notion, referred to as Coalitional Contingent Threats Equilibrium (CCTE), however, assumes that all deviating agreements are publicly observable. We show that UCPE is, generically, a refinement of Nash equilibrium. Unlike the case with CPE, this crucial containment property does not trivially follow from the definitions.

The blocking device proposed here is a "threat" which is a weakening of the notion of "trumping" used by Kahn and Mookherjee (1989) to define CPE and a strengthening of the notion of an "objection", used to define a CCTE. The key to arriving at our equilibrium is a "lateral induction" condition. Its intuition is derived from the idea of forwards induction, which plays a critical role in sequential games (see Kohlberg and Mertens (1986), Cho and Kreps (1987), Banks and Sobel (1987), etc.). It serves two critical roles. First, it bridges an informational asymmetry problem that arises in the absence of the nestedness assumption. This role is not entirely unexpected, given the links with forward induction. Second, it ensures that, generically, every UCPE is also a Nash equilibrium. This property is somewhat of a surprise since, a priori, the Nash best-response property is not a necessary condition for immunity from blocking when universal coalition formation is possible.

We shall briefly discuss some of the issues raised above that are peculiar to our problem before proceeding to the formal model.

(i) *Information asymmetry:* Universal coalition formation has been analyzed by Greenberg (1989, 1989a) in terms of "coalitional contingent threats" and "coalitional commitment". However, he assumes that all deviating agreements are made openly and are, therefore, public information. This assumption is, in general, not usual in the analysis of non-cooperative strategic-form games where only "equilibrium" strategies are assumed to be common knowledge. We permit deviating agreements to be made privately. Hence, if a new coalition forms which includes some players defecting from another coalition, the defectors have information which is unobservable to the other players joining the new coalition about the agreement they are defecting from. This is the root of the information asymmetry in our model. The absence of a resolution to the information problem had led to the nestedness assumption in Bernheim, *et al.* (1987).

(ii) *Stable partitions of the agreements space:* In the absence of a nestedness assumption, a definition of a stable agreement based on recursion on the number of players *a la* Bernheim, *et al.* (1987) cannot be given. We show that a characterization in terms of a partition of the space of agreements in the sense of von Neumann and Morgenstern (1947) abstract stable sets (which is the alternative proposed by Greenberg (1989)) also fails because of the existence of cycles in the ordering of agreements. Moreover, a characterization in terms of a weaker notion of semi-stability (as in Kahn and Mookherjee (1989)) is not possible, since a unique semi-stable partition of the space of agreements is not guaranteed. Given that we consider finite games, these difficulties are in sharp contrast with the proposition that a unique stable (and semi-stable) partition exists in environments satisfying the nestedness assumption (see Greenberg (1989) and Kahn and Mookherjee (1989)) and with the corresponding existence of a unique "labelling" system in characteristic function form games (Dutta *et al* (1989).

We derive a characterization of stable agreements using a notion of minimal semi-stability (Kahn and Mookherjee (1989)). The UCPE corresponds to n-player agreements that are threatened only by "strictly bad" agreements. A weaker notion would correspond to agreements that are not threatened by any "strictly good" agreements. We show that the latter, however, fails the Nash equilibrium test.

(iii) *Relationship with other equilibrium concepts:* The relationships among the blocking devices ("threat"; "trump", "objection") does not translate into a relationship among the corresponding equilibrium concepts (UCPE, CPE, CCTE). To check for stability of an agreement, we need to examine the entire hierarchy of blocking agreements in which the stability of each agreement depends on the stability of the agreement that blocks it. A weaker (or stronger) blocking device could make it both easier and more difficult to block a given n-player agreement, depending on the length of the hierarchy.

A priori, we should not expect our solution to be a refinement of Nash equilibrium, though this may appear paradoxical at first glance. In games with universal coalition formation, the Nash best-response property is not necessary for stability of an agreement in the von Neumann-Morgenstern sense. A Nash equilibrium corresponds to an n-player agreement that is not threatened by any one-player deviation. However, it is possible for an n-player agreement to be non-credibly threatened by a one-player deviation. The deviator may subsequently wish to form a coalition with other players and deviate from the initial deviation in a self-enforcing manner. Thus, the n-player agreement (which is non-Nash) is stable (in the von

Neumann-Morgenstern sense). This is a consequence of the fact that blocking coalitions are non-nested and no player can commit to an agreement.

In conjunction with the fact that we have (a) universal coalition formation and (b) public unobservability of deviating agreements, the Nash refinement property of our solution makes it more attractive than CPE and CCTE in characterizing stable agreements in non-cooperative games with pre-play non-binding communication within coalitions. If eventually the objective of this branch of game theory is to provide a framework that unifies cooperative and non-cooperative analysis, it is comforting to know that agreements that are stable (in a cooperative sense) under pre-play private communication within unrestricted coalitions do indeed satisfy the fundamental non-cooperative requirement of Nash equilibrium.

2. THE MODEL

$N = \{1, \ldots, n\}$ is a set of *players*. $\mathcal{H}$ is the *set of non-empty subsets of* N. For every $H \in \mathcal{H}$, $-H$ is the complement of H in N. Given $x = (x_i)_{i \in N} \in X$, for every $H \in \mathcal{H}$, let $x_H = (x_i)_{i \in H}$ and for all $i \in N$, $x_{-i} = (x_j)_{j \in N \setminus \{i\}}$. M is a (finite) set of *joint moves* by the members of N, with M_i being the set of moves for i. Every $i \in N$ has a *utility function* u_i: $M \to \mathbb{R}$. A *game* Γ is the triple $\langle N, M, (u)_{i \in N} \rangle$.

An *agreement* is a pair $[m, H] \in M \times \mathcal{H}$, with $\mathcal{A}$ denoting the set of all agreements.

An agreement is interpreted as a specification of (a) the moves to be taken by all parties to the agreement, given the moves of all other players, and (b) the set of players forming the agreement.

For any H, I, $J \in \mathcal{H}$ and m', $m'' \in M$, define $X[m'_H, m''_I; J] = \{\hat{m} \in M: \forall i \in J, u_i(m'_H, \hat{m}_{-H}) > u_i(m''_I, \hat{m}_{-I})\}$.

The set $X[m'_H, m''_I; J]$ contains all profiles of moves, $\hat{m}$, such that everybody in coalition J prefers a deviation by coalition H from $\hat{m}$ to m'_H to a deviation by coalition I from $\hat{m}$ to m''_I.

Define a binary relation $\succ$ on $\mathcal{A} \times \mathcal{A}$ as follows: $[\hat{m}, H] \succ [m', J]$ if $H \cap J \neq \emptyset$ and $\forall i \in H \backslash J$,

$$X[\hat{m}_H, m'_i; H] = X[\hat{m}_H, m'_i; H \cap J].$$

$X[\hat{m}_H, m'_i; H] \subseteq X[\hat{m}_H, m'_i; H \cap J]$ is immediate. The reverse containment requires that, given an initial profile m, if everybody common to both coalitions H and J prefers a deviation by coalition H to $\hat{m}_H$ to a deviation by a player i in H but not in J to m'_i, then everybody in H must display the same preference. The crucial role played by this relation is made clear after the next definition.

Define another binary relation $\succ\succ$ on $\mathcal{A} \times \mathcal{A}$ as follows: $[\hat{m}, H] \succ\succ [m', J]$ if

(i) for all $i \in H$, $u_i(\hat{m}) > u_i(m')$,

(ii) $\hat{m}_{-H} = m'_{-H}$, and

(iii) $[\hat{m}, H] \succ [m', J]$

If $[\hat{m}, H] \succ\succ [m', J]$, we say that $[\hat{m}, H]$ *threatens* $[m', J]$. Moreover,

$[\hat{m}, H]$ and $[m', J]$ are referred to, respectively, as the *source* and the *target* (of the threat). Members of the sets $H \cap J$ and $H \backslash J$ are referred to, respectively, as the *defectors* and the *recruits*.

$[\hat{m}, H]$ threatens $[m', J]$ if three conditions are satisfied. Both the defectors and the recruits must be made better off than they were in the target agreement (condition (i)). This provides the incentive for the formation of the source agreement. The moves of all players not party to the source agreement are held fixed at the values specified by the target agreement (condition (ii)).

Condition (iii) is a *lateral induction* condition. If the defectors know what the profile of moves in target agreement is, they must convince the recruits that it is indeed m'. (Subsequently, we shall argue that defectors will always be fully informed.) When the target agreement is a candidate "equilibrium", of course, we presume that all the moves would be common knowledge. However, all other agreements are out of equilibrium and each recruit would know only what he/she is playing. The moves to be played by the recruits and the defectors in the source agreement are observable to everybody party to that agreement. The defectors make the following implicit "speech" (see Cho and Kreps (1987) for other examples of such "speeches") to each recruit *i*.

> *"We want to convince you that the target agreement contains the profile* m'_{-i}. *The only situation in which we would have an incentive to defect from the target is when we stand to benefit from the defection. If the target contains an* $N\backslash\{i\}$*-profile of moves for which you would be made worse off by joining the source agreement, then at least one defector would be made worse off as well. If there is any* $N\backslash\{i\}$*-profile of moves that makes us better*

off from the defection, then you would be better off as well. This coincidence of interests is captured by the equivalence of the sets $X[\cdot; H]$ and $X[\cdot; H \cap J]$. $\hat{m}_H$ are the moves that you can observe if the defection occurs, and m'_i is the move that you can observe in the absence of the defection. Hence, the fact that we are proposing the formation of the source agreement must convey a signal that we are truthfully revealing our information about the moves in the target agreement, since we have no incentive to lie."

Begin with an n-player agreement that is common knowledge and identify the agreements that threaten it. The defectors are fully informed at this point. Next, identify the agreements that threaten the second agreement. Again the defectors are fully informed. And so on. By inductively applying the lateral induction argument above, we can ensure that defectors will always be fully informed about the moves profile in any target agreement in the hierarchy of agreements.

A stronger blocking device is used in the work of Bernheim, *et al.* (1987), Greenberg (1989) and Kahn and Mookherjee (1989). This can be expressed as a "trumping" relation (Kahn and Mookherjee (1989)):

Define a third binary relation $\succ\!\succ\!\succ$ on $\mathcal{A} \times \mathcal{A}$ as follows: $[\hat{m}, H] \succ\!\succ\!\succ [m', J]$ if

(i) $[\hat{m}, H] \succ\!\succ [m', J]$

(ii) $H \subseteq J$.

If $[\hat{m}, H] \succ\!\succ\!\succ [m', J]$, then $[\hat{m}, H]$ *trumps* $[m', J]$. The trumping relation requires that the threat relation be met and a nestedness assumption (i.e. condition (ii) above) be satisfied. Our definition appears somewhat different from the one given in Kahn and Mookherjee

(1989). The basic difference is the lateral induction condition which is a pre-requisite for a threat. However, note that the sets $X[\cdot; H]$ and $X[\cdot; H \cap J]$ are trivially equivalent in the case where $H \subseteq J$. Hence, the nestedness assumption ensures that this condition is automatically met.

3. THE GOOD, THE BAD AND THE UGLY

Our objective is to determine whether an agreement is stable, i.e. it is never threatened or is threatened only by an agreement that is threatened by an agreement that is never threatened or ...and so on. This logic follows the von Neumann and Morgenstern stable sets approach adopted by Greenberg (1989), Kahn and Mookherjee (1989) and Dutta *et al* (1989). To identify "stable" agreements, we must answer the following question: does there exist a unique "stable" (failing which a "semi-stable") partition of $\mathcal{A}$? These notions are defined as follows.

Consider three subsets of $\mathcal{A}$: the *good*, the *bad* and the *ugly*.

The *set of bad agreements*, $\mathcal{B}(\succ\succ)$, is defined by $\mathcal{B}(\succ\succ) = \{[m', H] \in \mathcal{A}: \exists[m'', J] \in \mathcal{G}(\succ\succ)$ such that $[m'', J] \succ\succ [m', H]\}$.

The *set of good agreements*, $\mathcal{G}(\succ\succ)$, is defined by $\mathcal{G}(\succ\succ) = \{[m', H] \in \mathcal{A}:$ if $\exists(m'', J] \in \mathcal{A}$ such that $[m'', J] \succ\succ [m', H]$, then $[m'', J] \in \mathcal{B}(\succ\succ)\}$.

The *set of ugly agreements*, $\mathcal{U}(\succ\succ)$, is defined as the complement of $\mathcal{G}(\succ\succ) \cup \mathcal{B}(\succ\succ)$ in $\mathcal{A}$.

$\mathcal{A}$ admits a $(\succ\succ)$-*stable* partition if $\mathcal{G}(\succ\succ) \cup \mathcal{B}(\succ\succ) = \mathcal{A}$ and $\mathcal{G}(\succ\succ) \cap \mathcal{B}(\succ\succ) = \emptyset$.

$\mathcal{A}$ admits a $(\succ\succ)$-*semi-stable* partition if $\mathcal{G}(\succ\succ) \cap \mathcal{B}(\succ\succ) = \emptyset$.

Correspondingly, we may define $\mathcal{G}(\succ\succ\succ)$ and $\mathcal{B}(\succ\succ\succ)$ and $\mathcal{U}(\succ\succ\succ)$ as the good, bad and ugly subsets of $\mathcal{A}$ by using trumping instead of a threat as a blocking device. Define $(\succ\succ\succ)$-stability and $(\succ\succ\succ)$-semi-stability analogously.

The Greenberg-Kahn-Mookherjee approach applies these notions to generate a solution concept in the following manner. First, show that there exists a unique $(\succ\succ\succ)$-stable partition of $\mathcal{A}$. Second, m is a solution to the game if $[m, N]$ is a good agreement.

PROPOSITION 0: (Greenberg (1989), Kahn and Mookherjee (1989)): *$\mathcal{A}$ admits a unique $(\succ\succ\succ)$-stable partition given that the underlying games are finite.*

This provides an alternative to the recursive approach of Bernheim, *et al.* (1987). Formally,

m is a *(nested) Coalition-proof Nash Equilibrium* (CPE) if $[m, N] \in \mathcal{G}(\succ\succ\succ)$.

The original motivation for a non-recursive definition was to facilitate extension of the fundamental idea of coalition-proofness to non-finite games. For our problem, even with finite games, the recursive approach cannot be used since we do not assume nestedness of deviating coalitions. Thus, the (semi-) stability approach appears to be the appropriate one for our purposes. However, the latter method would run into difficulties if a $(\succ\succ)$-(semi-)stable partition fails to exist or is not unique. Ideally, we would like to have a unique $(\succ\succ)$-stable partition. A $(\succ\succ)$-semi-stable partition that is not $(\succ\succ)$-stable contains an ugly set, which makes the solution concept somewhat ambiguous. Once we replace trumping with a threat as the blocking device, neither existence nor uniqueness of a stable partition is guaranteed, even in finite games. This is shown in the Appendix. Hence, we turn to an alternative partition of $\mathcal{A}$, called minimal semi-stability.

4. MINIMAL ($\succ\succ$)-SEMI-STABILITY

Consider the following construction from Kahn and Mookherjee (1989). Define the sets $\mathcal{G}_o^*$ and $\mathcal{B}_o^*$ as follows:

$\mathcal{G}_o^* = \{[m, H] \in \mathcal{A}$: $\exists$ no $[m', J] \in \mathcal{A}$ such that $[m', J] \succ\succ [m, H]\}$.

$\mathcal{B}_o^* = \{[m, H] \in \mathcal{A}$: $\exists [m', J] \in \mathcal{G}_o^*$ such that $[m', J] \succ\succ [m, H]\}$.

Next, inductively define $\mathcal{G}_z^*$, $\mathcal{B}_z^*$ with $z = 1, 2,\ldots$ as follows:

$\mathcal{G}_z^* = \{[m, H] \in \mathcal{A}$: if $[m', H'] \succ\succ [m, H]$, then $[m', H'] \in \mathcal{B}_{z-1}^*\}$.

$\mathcal{B}_z^* = \{[m, H] \in \mathcal{A}$: $\exists [m', H'] \in \mathcal{G}_z^*$ such that $[m', H'] \succ\succ [m, H]\}$.

For all z, $\mathcal{G}_{z-1}^* \subseteq \mathcal{G}_z^*$ and $\mathcal{B}_{z-1}^* \subseteq \mathcal{B}_z^*$, by the definitions given above. Define $\mathcal{G}^* = \bigcup_{z=0}^{\infty} \mathcal{G}_z^*$ and $\mathcal{B}^* = \bigcup_{z=0}^{\infty} \mathcal{B}_z^*$. Observe that if $[m, H] \in \mathcal{G}^*$ and there exists $[m', H'] \in \mathcal{A}$ such that $[m', H'] \succ\succ [m, H]$, then there exists z such that $[m, H] \in \mathcal{G}_z^*$ and $[m', H'] \in \mathcal{B}_{z-1}^*$. Hence, $[m', H'] \in \mathcal{B}^*$. Conversely, if $[m, H] \in \mathcal{B}^*$, then there exists z such that $[m, H] \in \mathcal{B}_z^*$ and $[m', H'] \in \mathcal{G}_z^*$ such that $[m', H'] \succ\succ [m, H]$. Hence, $[m', H'] \in \mathcal{G}^*$. $\mathcal{G}^*$ and $\mathcal{B}^*$ satisfy the definition of good and bad sets respectively.

In addition, we claim that $\mathcal{G}^* \cap \mathcal{B}^* = \emptyset$.

Suppose otherwise, i.e. $[m, H] \in \mathcal{G}^* \cap \mathcal{B}^*$. By construction, for some z, $[m, H] \in \mathcal{G}_z^* \cap \mathcal{B}_z^*$. $[m, H] \in \mathcal{B}_z^*$ implies that there exists $[m', H'] \in \mathcal{G}_z^*$ such that $[m', H'] \succ\succ [m, H]$. Since $[m, H] \in \mathcal{G}_z^*$ as well, $[m', H'] \in \mathcal{B}_{z-1}^*$.

Hence, there exists $[m'', H''] \in \mathcal{G}^*_{z-1}$ such that $[m'', H''] \succ\succ [m', H']$. Since $[m', H'] \in \mathcal{G}^*_z$, and $[m'', H''] \succ\succ [m', H']$, $[m'', H''] \in \mathcal{B}^*_{z-1}$. Repeating this argument, we conclude that there exists $[m, H] \in \mathcal{G}^*_o \cap \mathcal{B}^*_o$. This is in contradiction with the definitions of $\mathcal{G}^*_o$ and $\mathcal{B}^*_o$.

Finally, define $\mathcal{U}^*$ as the complement of $\mathcal{G}^* \cup \mathcal{B}^*$ in $\mathcal{A}$. Note that every agreement in $\mathcal{G}^*_o$ is a good agreement for any ($\succ\succ$)-semi-stable partition and, therefore, every agreement in $\mathcal{B}^*_o$ is a bad agreement for any ($\succ\succ$)-semi-stable partition. By applying an induction argument for all $\mathcal{G}^*_z$ and $\mathcal{B}^*_z$, we have just shown the existence of a partition of $\mathcal{A}$ which satisfies a property of "minimal" semi-stability, which is defined as follows.

A *minimal ($\succ\succ$)-semi-stable partition* $\{\mathcal{G}^*(\succ\succ), \mathcal{B}^*(\succ\succ), \mathcal{U}^*(\succ\succ)\}$ of $\mathcal{A}$ is one that satisfies: $\mathcal{G}^*(\succ\succ) \subseteq \mathcal{G}(\succ\succ)$ and $\mathcal{B}^*(\succ\succ) \subseteq \mathcal{B}(\succ\succ)$ for every ($\succ\succ$)-semi-stable partition $\{\mathcal{G}(\succ\succ), \mathcal{B}(\succ\succ), \mathcal{U}(\succ\succ)\}$ of $\mathcal{A}$. We refer to the elements of this partition as *strictly good*, *strictly bad* and *strictly ugly*.

Having settled on a unique method of partitioning the space of agreements, we now have a criterion for testing the stability of an n-player agreement.

m is a *Universal Coalition-proof Equilibrium* (UCPE) if $[m, N] \in \mathcal{G}^*(\succ\succ)$.

A weaker notion could also be defined as follows:

m is a *weak Universal Coalition-proof Equilibrium* (W-UCPE) if $[m, N] \in \mathcal{G}^*(\succ\succ) \cup \mathcal{U}^*(\succ\succ)$.

Our solution concept says that an n-player agreement is stable if it is threatened only by strictly bad agreements. The weaker concept says that an n-player agreement is stable if it is not threatened by any strictly good agreements. In the following section we shall argue that the latter is not a plausible notion of stability in a non-cooperative context.

5. EXAMPLES AND RELATIONSHIP WITH OTHER SOLUTION CONCEPTS

In this section, we shall explore the relationship of UCPE with other solution concepts that have been proposed to apply to the same class of problems. These include CPE, an equilibrium concept based on Coalitional Contingent Threats (Greenberg (1989, 1989a)) and Nash equilibrium. Clearly, the set of strong Nash equilibria (Aumann (1959)) are contained in the UCPE set. Hence, UCPE exist when the former exists. A general proof of existence of UCPE would face the same difficulties as those associated with CPE and other coalitional concepts.

First, we shall give some additional definitions.

5.1 ADDITIONAL DEFINITIONS

We consider an equilibrium concept based on the notion of Coalitional Contingent Threats, proposed by Greenberg (1989, 1989a) to characterize stable agreements when universal coalition formation is possible (i.e. precisely the environment that we have set about to study) under the assumption that all negotiations are made publicly.

Define a binary relation $\twoheadrightarrow$ on $\mathcal{A} \times \mathcal{A}$ as follows: $[\hat{m}, H] \twoheadrightarrow [m', J]$ if

(i) $\forall i \in H$, $u_i(\hat{m}) > u_i(m')$, and

(ii) $\hat{m}_{-H} = m'_{-H}$.

If $[\hat{m}, H] \twoheadrightarrow [m', J]$, we say that $[m, H]$ is an *objection* to $[m', J]$. Observe that an objection involves precisely the same requirements as a threat except for the lateral induction criterion. The latter is not used since negotiations are assumed to be made in public. The underlying process is as follows (see Greenberg (1989, 1989a)). An n-player agreement $[m, N]$ is on the table. A coalition J may openly declare that it objects to the agreement and will adopt m'_J instead provided the remaining players play m_{-J}. Another coalition, say H, can then object to the agreement $[m' \equiv (m'_J, m_{-J}), J]$ by threatening to play $\hat{m}_H$ provided the remaining players play m'_{-H}. This process continues until no coalition objects to a proposed profile, taking into account the possible reaction of other coalitions.

Define the sets $\mathcal{G}^*(\twoheadrightarrow)$, $\mathcal{B}^*(\twoheadrightarrow)$ and $\mathcal{U}^*(\twoheadrightarrow)$ such that $\{\mathcal{G}^*(\twoheadrightarrow), \mathcal{B}^*(\twoheadrightarrow), \mathcal{U}^*(\twoheadrightarrow)\}$ is a minimal $(\twoheadrightarrow)$-semi-stable partition of $\mathcal{A}$. The construction of this partition is analogous to that given for the relation $\succ\succ$.

m is a *Coalitional Contingent Threat Equilibrium* (*CCTE*) if $[m, N] \in \mathcal{G}^*(\twoheadrightarrow)$. m is a *weak Coalitional Contingent Threat Equilibrium* (*W-CCTE*) if $[m, N] \in \mathcal{G}^*(\twoheadrightarrow) \cup \mathcal{U}^*(\twoheadrightarrow)$.

Observe that a CCTE does not satisfy a fundamental property of Nash equilibrium, which is defined as follows:

m is a *Nash Equilibrium* if $\exists$ no $[m', \{i\}] \in \mathcal{A}$ such that $[m', \{i\}] \succ\succ\succ [m, N]$.

A priori, only CPE satisfies the Nash criterion by definition, since the condition given above follows from the condition that $[m, N] \in \mathcal{G}(\succ\succ\succ)$.

5.2 PROPOSITIONS

The points made by the examples below do not rely on any non-genericity in the payoffs. The first proposition shows that even though $\succ\succ$ is weaker than $\succ\succ\succ$, this does not translate into a relationship between the corresponding solution concepts.

PROPOSITION 1: *Neither UCPE nor W-UCPE is logically related to CPE.*

Proof: To show that there are no logical containment relationships between either UCPE or W-UCPE and CPE, consider the following examples.

EXAMPLE 1:

[Insert Figure 1 here]

Consider the game given above. Player 1 chooses from $\{T, B\}$, 2 chooses from $\{L, R\}$ and 3 chooses from $\{\ell, r\}$. This example will show that $[(B, R, \ell), \{1, 2, 3\}] \in \mathcal{B}(\succ\succ\succ)$ and $[(B, R, \ell), \{1, 2, 3\}] \in \mathcal{G}^*(\succ\succ)$.

A Nash equilibrium of the game is (B, R, ℓ). It is not a CPE since $[(T, L, \ell), \{1, 2\}] \succ\succ\succ [(B, R, \ell), \{1, 2, 3\}]$. Check that $[(T, L, \ell), \{1, 2\}] \in \mathcal{G}(\succ\succ\succ)$. We claim, however, that $[(B, R, \ell), \{1, 2, 3\}] \in \mathcal{G}^*(\succ\succ)$. The argument is in several steps.

(I) $[(T, L, \ell), \{1, 2\}] \succ\succ\succ [(B, R, \ell), \{1, 2, 3\}]$ implies that $[(T, L, \ell), \{1, 2\}] \succ\succ [(B, R, \ell), \{1, 2, 3\}]$.

(II) Next, we claim that $[(B, L, r), \{1, 3\}] \succ\succ [(T, L, \ell), \{1, 2\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(B, r), \ell; \{1, 3\}] = \{(B, L, \ell), (B, L, r), (T, L, \ell), (T, L, r)\} = X[(B, r), \ell; \{1\}]$. Therefore, $[(B, L, r), \{1, 3\}] \succ [(T, L,$

ℓ), {1, 2}].

Since [(B, L, r), {1, 3}] is not threatened by any other agreement, it is in $\mathcal{G}^*(\succ\succ)$. Hence, [($T$, L, ℓ), {1, 2}] $\in \mathcal{B}^*(\succ\succ)$. Since [($T$, L, ℓ), {1, 2}] is the unique agreement that threatens [(B, R, ℓ), {1, 2, 3}], the latter is in $\mathcal{G}^*(\succ\succ)$.

EXAMPLE 2:

[Insert Figure 2 here]

This game has player 3 choosing from {ℓ, c, r}. The remaining players' strategy spaces are the same as in the previous game. This example will show that [(T, L, c), {1, 2, 3}] $\in \mathcal{G}(\succ\succ\succ)$ and [(T, L, c), {1, 2, 3}] $\in \mathcal{B}^*(\succ\succ)$.

Consider the Nash equilibrium (T, L, c). It is also a CPE. This may be checked as follows. The only agreement that trumps [(T, L, c), {1, 2, 3}} is [(B, R, ℓ), {1, 2, 3}]. However, [(T, L, ℓ), {1, 2}] $\succ\succ\succ$ [(B, R, ℓ), {1, 2, 3}]. [(T, L, ℓ), {1, 2}] $\in \mathcal{G}(\succ\succ\succ)$ since it is not trumped by any agreements.

We claim that [(T, L, c), {1, 2, 3}] $\in \mathcal{B}^*(\succ\succ)$. The argument is as follows:

(I) [(T, L, ℓ), {1, 2}] $\succ\succ\succ$ [(B, R, ℓ), {1, 2, 3}] implies that [(T, L, ℓ), {1, 2}] $\succ\succ$ [(B, R, ℓ), {1, 2, 3}].

(II) Next, we claim that [(B, L, r), {1, 3}] $\succ\succ$ [(T, L, ℓ), {1, 2}]. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that X[(B, r), ℓ; {1, 3}] = {(B, L, ℓ), (B, L, r), (T, L, ℓ), (T, L, r), (B, L, c), (T, L, c)} = X[(B, r), ℓ; {1}]. Therefore, [(B, L, r), {1, 3}] $\succ$ [(T, L, ℓ), {1, 2}].

Since $[(B, L, r), \{1, 3\}]$ is not threatened by any other agreement, it is in $\mathcal{G}^*(\succ\succ)$. Hence, $[(T, L, \ell), \{1, 2\}] \in \mathcal{B}^*(\succ\succ)$. Since $[(T, L, \ell), \{1, 2\}]$ is the only agreement that threatens $[(B, R, \ell), \{1, 2, 3\}]$, the latter is in $\mathcal{G}^*(\succ\succ)$. $[(B, R, \ell), \{1, 2, 3\}] \succ\succ [(T, L, c), \{1, 2, 3\}]$ implies that $[(T, L, c), \{1, 2, 3\}] \in \mathcal{B}^*(\succ\succ)$. ∎

The second proposition shows that even though $\succ\succ$ is stronger than $\gg$, this does not translate into a relationship between the corresponding solution concepts.

PROPOSITION 2: *Neither UCPE nor W-UCPE is logically related to either CCTE or W-CCTE.*

Proof: Consider the following examples.

EXAMPLE 3:

[Insert Figure 3 here]

Consider the game given above. Player 1 chooses from $\{T, B\}$, player 2 chooses from $\{L, C, R\}$ and player 3 chooses from $\{\ell, r\}$. The example shows that $[(B, C, \ell), \{1, 2, 3\}] \in \mathcal{G}^*(\succ\succ)$ and $[(B, C, \ell), \{1, 2, 3\}] \in \mathcal{B}^*(\gg)$.

First, we show that $[(B, C, \ell), \{1, 2, 3\}] \in \mathcal{G}^*(\succ\succ)$. The argument is in several steps:

(I) $[(T, L, \ell), \{1, 2\}] \succ\succ\succ [(B, C, \ell), \{1, 2, 3\}]$, which implies that $[(T, L, \ell), \{1, 2\}] \succ\succ [(B, C, \ell), \{1, 2, 3\}]$.

(II) We claim that $[(B, L, r), \{1, 3\}] \succ\succ [(T, L, \ell), \{1, 2\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(B, r), \ell; \{1, 3\}] = \{(B, L, \ell), (B, L, r), (T, L, \ell), (T, L, r)\} = X[(B, r), \ell; \{1\}]$. Therefore, $[(B, L, r), \{1, 3\}] \succ [(T, L, \ell), \{1, 2\}]$.

Since $[(B, L, r), \{1, 3\}]$ is not threatened by any other agreement, it is in $\mathcal{G}^*(\succ\succ)$. Hence, $[(T, L, \ell), \{1, 2\}] \in \mathcal{B}^*(\succ\succ)$. Since $[(T, L, \ell), \{1, 2\}]$ is the only agreement that threatens $[(B, C, \ell), \{1, 2, 3\}]$, the latter is in $\mathcal{G}^*(\succ\succ)$.

Next, we show that $[(B, C, \ell), \{1, 2, 3\}] \in \mathcal{B}^*(\twoheadrightarrow)$. First, observe that $[(T, L, \ell), \{1, 2\}] \succ\succ [(B, C, \ell), \{1, 2, 3\}]$ implies that $[(T, L, \ell), \{1, 2\}] \twoheadrightarrow [(B, C, \ell), \{1, 2, 3\}]$. Moreover, $[(B, L, r), \{1, 3\}] \succ\succ [(T, L, \ell), \{1, 2\}]$ implies that $[(B, L, r), \{1, 3\}] \twoheadrightarrow [(T, L, \ell), \{1, 2\}]$. Finally, observe that $[(B, R, r), \{2\}] \twoheadrightarrow [(B, L, r), \{1, 3\}]$.

Since there is no agreement that constitutes an objection to $[(B, R, r), \{2\}]$, the latter is in $\mathcal{G}^*(\twoheadrightarrow)$. $[(B, R, r), \{2\}] \twoheadrightarrow [(B, L, r), \{1, 3\}]$ implies that $[(B, L, r), \{1, 3\}] \in \mathcal{B}^*(\twoheadrightarrow)$. Since $[(B, L, r), \{1, 3\}]$ is the only agreement that is an objection to $[(T, L, \ell), \{1, 2\}]$, the latter is in $\mathcal{G}^*(\twoheadrightarrow)$. $[(T, L, \ell), \{1, 2\}] \twoheadrightarrow [(B, C, \ell), \{1, 2, 3\}]$ implies that $[(B, C, \ell), \{1, 2, 3\}] \in \mathcal{B}^*(\twoheadrightarrow)$.

Next, consider the following example.

EXAMPLE 4:

[Insert Figure 4 here]

In the game above, Player 1 chooses from $\{T, B\}$ and player 2 chooses from $\{L, R\}$. This example shows that $[(T, L), \{1, 2\}] \in \mathcal{G}^*(\twoheadrightarrow)$ and $[(T, L), \{1, 2\}] \in \mathcal{B}^*(\succ\succ)$.

$[(T, R), \{2\}] \twoheadrightarrow [(T, L), \{1, 2\}]$ and $[(B, R), \{1\}] \twoheadrightarrow [(T, R), \{2\}]$. Also, $[(B, L), \{1\}] \twoheadrightarrow [(T, L), \{1, 2\}]$ and $[(B, R), \{2\}] \twoheadrightarrow [(B, L), \{1\}]$. Since there is no agreement that constitutes an objection to $[(B, R), \{1\}]$, the latter is in $\mathcal{G}^*(\twoheadrightarrow)$. Hence, $[(B, R), \{1\}] \twoheadrightarrow [(T, R), \{2\}]$ implies that $[(T, R), \{2\}] \in \mathcal{B}^*(\twoheadrightarrow)$. Moreover, since there is no agreement that constitutes an objection to $[(B, R), \{2\}]$, the latter is in $\mathcal{G}^*(\twoheadrightarrow)$. Hence,

$[(B, R), \{2\}] \gg [(B, L), \{1\}]$ implies that $[(B, L), \{1\}] \in \mathcal{B}^*(\gg)$. Given that the set of agreements that are objections to $[(T, L), \{1, 2\}]$ is $\{[(T, R), \{2\}], [(B, L), \{1\}]\} \subseteq \mathcal{B}^*(\gg)$, we have $[(T, L), \{1, 2\}] \in \mathcal{G}^*(\gg)$.

Next, observe that $[(T, R), \{2\}] \succ\succ [(T, L), \{1, 2\}]$ and $[(B, L), \{1\}] \succ\succ [(T, L), \{1, 2\}]$. However, there is no agreement that threatens either $[(T, R), \{2\}]$ or $[(B, L), \{1\}]$. Hence, $\{[(T, R), \{2\}], [(B, L), \{1\}]\} \subseteq \mathcal{G}^*(\succ\succ)$, which implies that $[(T, L), \{1, 2\}] \in \mathcal{B}^*(\succ\succ)$ since the set of agreements that threaten it is $\{[(T, R), \{2\}], [(B, L), \{1\}]\}$. ■

The last example also shows that a CCTE need not be a Nash equilibrium. This may be expected of a UCPE as well. However, we have the following result.

PROPOSITION 3: *Generically, a UCPE is also a Nash equilibrium.*

<u>Proof:</u> Suppose m is a UCPE and is not a Nash equilibrium. Then there exists $i \in N$ such that $[(m_i', m_{-i}), \{i\}] \succ\succ\succ [m, N]$ and, therefore, $[(m_i', m_{-i}), \{i\}] \succ\succ [m, N]$. By definition of UCPE, $[m, N] \in \mathcal{G}^*(\succ\succ)$. Hence, $[(m_i', m_{-i}), \{i\}] \in \mathcal{B}^*(\succ\succ)$ and, therefore, there must exist $H \in \mathcal{H}$ and an agreement $[m'', \{\{i\} \cup H\}] \in \mathcal{G}^*(\succ\succ)$ such that $[m'', \{\{i\} \cup H\}] \succ\succ [(m_i', m_{-i}), \{i\}]$. We shall consider two cases:

<u>Case (i):</u> There exists $j \in H$ such that $u_j(m) > u_j(m''_{H\cup(i)}, m_{-(H\cup(i))})$ By the lateral induction requirement, i.e. condition (iii) of the definition of a threat, we must have $X[m''_{H\cup(i)}, m_j; H\cup\{i\}] = X[m''_{H\cup(i)}, m_j; \{i\}]$, i.e. $u_i(m) > u_i(m''_{H\cup(i)}, m_{-(H\cup(i))})$. Recall that $u_i(m''_{H\cup(i)}, m_{-(H\cup(i))}) = u_i(m'')$. We have a contradiction with the requirement that $[m'', \{\{i\} \cup H\}] \succ\succ [(m_i', m_{-i}), \{i\}] \succ\succ [m, N]$ must imply $u_i(m'') > u_i(m_i', m_{-i}) > u_i(m)$.

<u>Case (ii):</u> There exists no $j \in H$ such that $u_j(m) > u_j(m''_{H\cup(i)}, m_{-(H\cup(i))})$

By the genericity assumption, $u_j(m) < u_j(m''_{H\cup\{i\}}, m_{-(H\cup\{i\})})$ for all $j \in H$. $[m'', \{\{i\} \cup H\}] \succ\succ [(m'_i, m_{-i}), \{i\}] \succ\succ [m, N]$ implies that $u_i(m''_{H\cup\{i\}}, m_{-(H\cup\{i\})}) = u_i(m'') > u_i(m)$. Given that for all $j \in H$, $u_j(m''_{H\cup\{i\}}, m_{-(H\cup\{i\})}) = u_j(m'') > u_j(m)$, we have $[m'', \{\{i\} \cup H\}] \in \mathcal{G}^*(\succ\succ)$ such that $[m'', \{\{i\} \cup H\}] \succ\succ\succ [m, N]$. This implies that $[m'', \{\{i\} \cup H\}] \succ\succ [m, N]$ and is in contradiction with the assumption that $[m, N] \in \mathcal{G}^*(\succ\succ)$. ■

The weaker notion W-UCPE does not, however, have the Nash equilibrium property. This can be seen from the following proposition. This makes the W-UCPE less desirable as a characterization of stability in a non-cooperative context.

PROPOSITION 4: *There is no logical relationship between W-UCPE and Nash equilibrium.*

Proof: From example 2, we know that a Nash equilibrium is not necessarily W-UCPE. To see that a W-UCPE is not necessarily a Nash equilibrium, consider the following example.

Example 5:

[Insert Figure 5 here]

Consider the game above. Player 1 chooses from $\{T, B\}$, 2 chooses from $\{L, R\}$ and 3 chooses from $\{\ell, r\}$. We claim that $[(T, L, \ell), \{1, 2, 3\}] \in \mathcal{U}^*(\succ\succ)$ whereas (T, L, ℓ) is not a Nash equilibrium. We need the following lemmata.

LEMMA 1: $\{[(B, R, \ell), \{1, 2\}], [(B, L, r), \{2, 3\}], [(T, L, \ell), \{1, 3\}]\} \subseteq \mathcal{U}^*(\succ\succ)$.

Proof of Lemma 1: We shall show that the relation $\succ\succ$ induces a cycle in the set of agreements $\{[(B, R, \ell), \{1, 2\}], [(B, L, r), \{2, 3\}], [(T, L, \ell), \{1, 3\}]\}$. The argument proceeds in several steps.

(I) We claim that $[(B, L, r), \{2, 3\}] \succ\succ [(B, R, \ell), \{1, 2\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(L, r), \ell; \{2, 3\}] = \{(B, L, \ell), (B, L, r), (B, R, \ell), (B, R, r)\} = X[(L, r), \ell; \{2\}]$. Therefore, $[(B, L, r), \{2, 3\}] \succ [(B, R, \ell), \{1, 2\}]$.

(II) Next, we claim that $[(T, L, \ell), \{1, 3\}] \succ\succ [(B, L, r), \{2, 3\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(T, \ell), B; \{1, 3\}] = \{(T, L, \ell), (B, L, \ell), (T, L, r), (B, L, r)\} = X[(T, \ell), B; \{3\}]$. Therefore, $[(T, L, \ell), \{1, 3\}] \succ [(B, L, r), \{2, 3\}]$.

(III) Finally, we claim that $[(B, R, \ell), \{1, 2\}] \succ\succ [(T, L, \ell), \{1, 3\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(B, R), L; \{1, 2\}] = \{(T, L, \ell), (B, L, \ell), (T, R, \ell), (B, R, \ell)\} = X[(B, R), L; \{1\})$. Therefore, $[(B, R, \ell), \{1, 2\}] \succ [(T, L, \ell), \{1, 3\}]$.

The steps (I)-(III) have generated a cycle of threatened agreements since $[(B, R, \ell), \{1, 2\}] \succ\succ [(T, L, \ell), \{1, 3\}] \succ\succ [(B, L, r), \{2, 3\}] \succ\succ [(B, R, \ell), \{1, 2\}] \succ\succ \ldots$

We claim that $[(B, R, \ell), \{1, 2\}] \in \mathcal{U}^*(\succ\succ)$. Suppose otherwise. There are two cases to be examined.

Suppose $[(B, R, \ell), \{1, 2\}] \in \mathcal{B}^*(\succ\succ)$, in which case it is threatened by an agreement in $\mathcal{G}^*(\succ\succ)$. However, there is no agreement other than $[(B, L, r), \{2, 3\}]$ that threatens $[(B, R, \ell), \{1, 2\}]$. Hence, $[(B, L, r), \{2, 3\}] \in \mathcal{G}^*(\succ\succ)$. Then $[(T, L, \ell), \{1, 3\}] \in \mathcal{B}^*(\succ\succ)$ since $[(T, L, \ell), \{1, 3\}]$

$\succ\succ [(B, L, r), \{2, 3\}]$. If $[(T, L, \ell), \{1, 3\}] \in \mathcal{B}^*(\succ\succ)$, we must have $[(B, R, \ell), \{1, 2\}] \in \mathcal{G}^*(\succ\succ)$ since there is no agreement other than $[(B, R, \ell), \{1, 2\}]$ that threatens $[(T, L, \ell), \{1, 3\}]$. Hence, we have a contradiction.

Suppose that $[(B, R, \ell), \{1, 2\}] \in \mathcal{G}^*(\succ\succ)$, in which case $[(B, L, r), \{2, 3\}] \in \mathcal{B}^*(\succ\succ)$, since $[(B, L, r), \{2, 3\}] \succ\succ [(B, R, \ell), \{1, 2\}]$. Thus, $[(T, L, \ell), \{1, 3\}] \in \mathcal{G}^*(\succ\succ)$ since $[(T, L, \ell), \{1, 3\}]$ is the only agreement that threatens $[(B, L, r), \{2, 3\}]$. If $[(T, L, \ell), \{1, 3\}] \in \mathcal{G}^*(\succ\succ)$, we must have $[(B, R, \ell), \{1, 2\}] \in \mathcal{B}^*(\succ\succ)$ since $[(B, R, \ell), \{1, 2\}] \succ\succ [(T, L, \ell), \{1, 3\}]$. Hence, we have a contradiction.

An analogous argument can be given to show that $[(T, L, \ell), \{1, 3\}] \in \mathcal{U}^*(\succ\succ)$ and $[(B, L, r), \{2, 3\}] \in \mathcal{U}^*(\succ\succ)$. ∎

LEMMA 2: $[(T, R, \ell), \{2\}] \in \mathcal{U}^*(\succ\succ)$.

<u>Proof of Lemma 2:</u> First, we claim that $[(B, R, \ell), \{1, 2\}] \succ\succ [(T, R, \ell), \{2\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(B, R), T; \{1, 2\}] = \{(T, L, \ell), (B, L, \ell), (T, R, \ell), (B, R, \ell), (T, R, r), (B, R, r)\} = X[(B, R), T; \{2\}]$. Therefore, $[(B, R, \ell), \{1, 2\}] \succ [(T, R, \ell), \{2\}]$.

Observe that neither $[(B, L, r), \{1, 2, 3\}]$ nor $[(B, L, r), \{2, 3\}]$ threatens $[(T, R, \ell), \{2\}]$. In each case, though conditions (i) and (ii) for a threat are met, condition (iii) is not satisfied. The unique agreement that threatens $[(T, R, \ell), \{2\}]$ is $[(B, R, \ell), \{1, 2\}]$. $[(B, R, \ell), \{1, 2\}] \in \mathcal{U}^*(\succ\succ)$ (from Lemma 1) implies that $[(T, R, \ell), \{2\}] \in \mathcal{U}^*(\succ\succ)$. ∎

To check that $[(T, L, \ell), \{1, 2, 3\}] \in \mathcal{U}^*(\succ\succ)$, observe that the only agreements that threaten it are $[(B, R, \ell), \{1, 2\}]$ and $[(T, R, \ell), \{2\}]$.

By Lemma 1 and Lemma 2, $\{[(B, R, \ell), \{1, 2\}], [(T, R, \ell), \{2\}]\} \subseteq \mathcal{U}^*(\succ\succ)$. ■

Since existence is always an issue with coalitional solution concepts, it is useful to know of a sufficient condition for a Nash containment property. Thus, we would have a concept that is weaker than UCPE and is also a Nash equilibrium refinement. Define an intermediate concept that is stronger than W-UCPE and weaker than UCPE as follows: m is a *quasi-weak Universal Coalition-proof Equilibrium* (Q-UCPE) if m is a W-UCPE and satisfies:

$[\exists[m', \{i\}] \in \mathcal{A}$ such that $[m', \{i\}] \succ\succ [m, N]] \Rightarrow [\exists[m'', \{j\}] \in \mathcal{A}$ such that $[m'', \{j\}] \succ\succ [m, N]$ and $[m'', \{j\}] \in \mathcal{B}^*(\succ\succ)]$.

The proof of Proposition 3 yields the following conclusion as well.

COROLLARY TO PROPOSITION 3: *Generically, a Q-UCPE is also a Nash equilibrium.*

7. CONCLUDING REMARKS

(i) A subgame-perfect version of UCPE can also be given. For a multi-stage game Γ, let γ be a proper subgame of Γ. An agreement is *perfectly good* if it is good in every subgame. An agreement is *perfectly bad* if it is bad in some subgame. All other agreements are *perfectly ugly*. Define a minimal perfect $(\succ\succ)$-semi-stable partition as in the earlier sections. All the propositions discussed above can be transported to this framework and by replacing CPE, CCTE and Nash equilibrium with the corresponding concepts requiring subgame-perfection.

(ii) We also apply the notion of lateral induction in a cooperative context in Chakravorti (1990) in a new definition of a Bargaining Set.

(iii) An interesting comparison is with the consistent Bargaining Set of Dutta *et al* (1989). Due to a fundamental restriction in the definition of the Bargaining Set (see Chakravorti (1990)), they avoid cycles of the form given in the Appendix of this paper. We have shown that despite the presence of such cycles in the ordering of agreements, an appealing characterization of stable agreements can be given. In particular, we have criteria that provide a justification for focusing on a subset of Nash equilibrium despite the fact that in environments with universal coalition formation possibilities, Nash equilibrium is not a necessary condition for (von Neumann and Morgenstern) stability of an outcome.

References

AUMANN, R. (1959): "Acceptable Points in General Cooperative *n*-Person Games," *Ann. Math. Stud.,* 40, 287-324.

AUMANN, R. AND M. MASCHLER (1964): "The Bargaining Set for Cooperative Games," Dresher, Shapley and Tucker (ed.) *Advances in Game Theory,* Princeton, NJ, Princeton University Press.

BANKS, J. AND J. SOBEL (1987): "Equilibrium Selection in Signaling Games," *Econometrica,* 55, 647-662.

BERNHEIM, D., B. PELEG AND M. WHINSTON (1987): "Coalition-proof Nash Equilibria, I: Concepts," *J. Econ. Theory,* 42, 1-12.

CHAKRAVORTI, B. (1990): "A $(\text{Consistent})^2$ Bargaining Set: Non-Cooperative Renegotiation in Cooperative Games," mimeo, University of Illinois at Urbana-Champaign.

CHO, I. AND D. KREPS (1987): "Signaling Games and Stable Equilibria," *Quart. J. Econ.*, 102, 179-221.

DUTTA, B., D. RAY, K. SENGUPTA AND R. VOHRA (1989): "A Consistent Bargaining Set," *J. Econ. Theory,* 49, 93-112.

GREENBERG, J. (1989): "Deriving Strong and Coalition-proof Nash Equilibrium from an Abstract System," *J. Econ. Theory,* 49, 195-202.

GREENBERG, J. (1989a): *The Theory of Social Situations,* Cambridge, UK, Cambridge University Press.

KAHN, C. AND D. MOOKHERJEE (1989): "The Good, the Bad and the Ugly: Coalition-proof Equilibrium in Games with Infinite Strategy Spaces," mimeo University of Illinois at Urbana-Champaign, *Games and Econ. Behav.* (forthcoming).

KOHLBERG, E. AND J.-F. MERTENS (1986): "On the Strategic Stability of

Equilibria," *Econometrica*, 54, 1003-38.

MAS-COLELL, A. (1989): "An Equivalence Theorem for a Bargaining Set," *J. Math. Econ.*, 18, 129-138.

SHAPLEY, L. AND M. SHUBIK (1984): in Shubik, *Game Theory in the Social Sciences,* Cambridge, MA, MIT Press.

VON NEUMANN, J. AND O. MORGENSTERN (1947): *Theory of Games and Economic Behavior,* Princeton, NJ, Princeton University Press.

APPENDIX

NON-EXISTENCE OF A UNIQUE ($\succ\succ$)-(SEMI)-STABLE PARTITION

PROPOSITION A.1 : *In general, $\mathcal{A}$ does not admit a ($\succ\succ$)-stable partition.*

Proof: The proof is by way of an example. We shall present a game for which $\mathcal{U}(\succ\succ) \neq \varnothing$.

Example 6: Consider the following game. Player 1 chooses from $\{T, B\}$, 2 chooses from $\{L, R\}$ and 3 chooses from $\{\ell, r\}$.

[Insert Figure 6 here]

We shall show that the relation $\succ\succ$ induces a cycle in the set of agreements $\{[(B, R, \ell), \{1, 2\}], [(B, L, r), \{2, 3\}], [(T, L, \ell), \{1, 3\}]\}$. The argument proceeds in several steps.

(I) We claim that $[(B, L, r), \{2, 3\}] \succ\succ [(B, R, \ell), \{1, 2\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(L, r), \ell; \{2, 3\}] = \{(B, L, \ell), (B, L, r), (B, R, \ell), (B, R, r)\} = X[(L, r), \ell; \{2\}]$. Therefore, $[(B, L, r), \{2, 3\}] \succ [(B, R, \ell), \{1, 2\}]$.

(II) We claim that $[(T, L, \ell), \{1, 3\}] \succ\succ [(B, L, r), \{2, 3\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(T, \ell), B; \{1, 3\}] = \{(T, L, \ell), (B, L, \ell), (T, L, r), (B, L, r)\} = X[(T, \ell), B; \{3\}]$. Therefore, $[(T, L, \ell), \{1, 3\}] \succ [(B, L, r), \{2, 3\}]$.

(III) We claim that $[(B, R, \ell), \{1, 2\}] \succ\succ [(T, L, \ell), \{1, 3\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(B, R), L; \{1, 2\}] = \{(T, L, \ell), (B, L, \ell), (T, R, \ell), (B, R, \ell)\} = X[(B, R), L; \{1\})$. Therefore, $[(B, R, \ell), \{1, 2\}] \succ [(T, L, \ell), \{1, 3\}]$.

The steps (I)-(III) have generated a cycle of threatened agreements since $[(B, R, \ell), \{1, 2\}] \succ\succ [(T, L, \ell), \{1, 3\}] \succ\succ [(B, L, r), \{2, 3\}] \succ\succ [(B, R, \ell), \{1, 2\}] \succ\succ \ldots$

We claim that $[(B, R, \ell), \{1, 2\}] \in \mathcal{U}(\succ\succ)$. Suppose otherwise. There are two cases to be examined.

Suppose $[(B, R, \ell), \{1, 2\}] \in \mathcal{B}(\succ\succ)$, in which case it is threatened by an agreement in $\mathcal{G}(\succ\succ)$. However, there is no agreement other than $[(B, L, r), \{2, 3\}]$ that threatens $[(B, R, \ell), \{1, 2\}]$. Hence, $[(B, L, r), \{2, 3\}] \in \mathcal{G}(\succ\succ)$. Then $[(T, L, \ell), \{1, 3\}] \in \mathcal{B}(\succ\succ)$ since $[(T, L, \ell), \{1, 3\}] \succ\succ [(B, L, r), \{2, 3\}]$. If $[(T, L, \ell), \{1, 3\}] \in \mathcal{B}(\succ\succ)$, we must have $[(B, R, \ell), \{1, 2\}] \in \mathcal{G}(\succ\succ)$ since there is no agreement other than $[(B, R, \ell), \{1, 2\}]$ that threatens $[(T, L, \ell), \{1, 3\}]$. Hence, we have a contradiction.

Suppose that $[(B, R, \ell), \{1, 2\}] \in \mathcal{G}(\succ\succ)$, in which case $[(B, L, r), \{2, 3\}] \in \mathcal{B}(\succ\succ)$, since $[(B, L, r), \{2, 3\}] \succ\succ [(B, R, \ell), \{1, 2\}]$. Thus, $[(T, L, \ell), \{1, 3\}] \in \mathcal{G}(\succ\succ)$ since $[(T, L, \ell), \{1, 3\}]$ is the only agreement that threatens $[(B, L, r), \{2, 3\}]$. If $[(T, L, \ell), \{1, 3\}] \in \mathcal{G}(\succ\succ)$, we must have $[(B, R, \ell), \{1, 2\}] \in \mathcal{B}(\succ\succ)$ since $[(B, R, \ell), \{1, 2\}] \succ\succ [(T, L, \ell), \{1, 3\}]$. Hence, we have a contradiction. ■

PROPOSITION A.2: *In general, $\mathcal{A}$ does not admit a unique $(\succ\succ)$-semi-stable partition.*

Proof: The proof is by way of an example. We shall present a game such that a cycle is generated as in the previous example. Each agreement in the cycle can be defined as both good and bad relative to corresponding re-definitions of the agreements in the cycle that threaten it.

Example 7: Consider the following game. Player 1 chooses from $\{T, B\}$, 2 chooses from $\{L, R\}$, 3 chooses from $\{\ell, r\}$ and 4 chooses from $\{U, D\}$.

[Insert Figure 7 here]

We shall show that the relation $\succ\!\succ$ induces a cycle of the form discussed above in the set $\{[(B, R, \ell, U), \{1, 2\}], [(B, L, r, U), \{2, 3\}], [(B, L, r, D), \{3, 4\}], [(T, L, \ell, U), \{1, 3, 4\}]$. The argument proceeds in several steps.

(I) We claim that $[(B, L, r, U), \{2, 3\}] \succ\!\succ [(B, R, \ell, U), \{1, 2\}]$. Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(L, r), \ell; \{2, 3\}] = \{(B, L, \ell, U), (B, L, r, U), (B, R, \ell, U), (B, R, r, U)\} = X[(L, r), \ell); \{2\})$. Therefore, $[(B, L, r, U), \{2, 3\}] \succ [(B, R, \ell, U), \{1, 2\}]$.

(II) We claim that $[(B, L, r, D), \{3, 4\}] \succ\!\succ [(B, L, r, U), \{2, 3\}]$, Conditions (i) and (ii) for a threat met. To check for condition (iii), verify that $X[(r, D), U; \{3, 4\}] = \{\{(B, L, \ell, U), (B, L, r, U), (B, L, \ell, D), (B, L, r, D)\} = X[(r, D), U; \{3\})$. Therefore, $[(B, L, r, D), \{3, 4\}] \succ [(B, L, r, U), \{2, 3\}]$.

(III) We claim that $[(T, L, \ell, U), \{1, 3, 4\}] \succ\!\succ [(B, L, r, D), \{3, 4\}]$, Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(T, \ell, U), B; \{1, 3, 4\}] = \{(T, L, \ell, U), (B, L, \ell, U), (T, L, r, U), (B, L, r, U), (T, L, r, D), (B, L, r, D), (T, L, \ell, D), (B, L, \ell, D)\} = X[(T, \ell, U), B; \{3, 4\})$. Therefore, $[(T, L, \ell, U), \{1, 3, 4\}] \succ [(B, L, r, D), \{3, 4\}]$.

(IV) We claim that $[(B, R, \ell, U), \{1, 2\}] \succ\!\succ [(T, L, \ell, U), \{1, 3, 4\}]$, Conditions (i) and (ii) for a threat are met. To check for condition (iii), verify that $X[(B, R), L; \{1, 2\}] = \{(T, L, \ell, U), (B, L, \ell, U), (T, R, \ell, U), (B, R, \ell, U), (B, L, \ell, D), (B, R, \ell, D)\} = X[(B, R), L; \{1\})$. Therefore, $[(B, R, \ell, U), \{1, 2\}] \succ [(T, L, \ell, U), \{1, 3, 4\}]$.

The steps (I)-(IV) have generated a cycle of threatened agreements since $((B, R, \ell, U), \{1, 2\}) \succ\succ ((T, L, \ell, U), \{1, 3, 4\}) \succ\succ ((B, L, r, D), \{3, 4\}) \succ\succ ((B, L, r, U), \{2, 3\}) \succ\succ ((B, R, \ell, U), \{1, 2\}) \succ\succ \ldots$

Also, check that for each one of the agreements in this cycle, there is only one agreement that threatens it.

The cycle generated above has the following structure:

[Insert Figure 8 here]

{**A**, **B**, **C**, **D**} is such that **A** $\succ\succ$ **B** $\succ\succ$ **C** $\succ\succ$ **D** $\succ\succ$ **A** $\succ\succ$.... **B** is the only agreement that threatens **A**. **C** is the only agreement that threatens **B**. **D** is the only agreement that threatens **C** and **A** is the only agreement that threatens **D**.

Suppose that **A** $\in \mathcal{G}(\succ\succ)$. Then **D** $\in \mathcal{B}(\succ\succ)$, **C** $\in \mathcal{G}(\succ\succ)$ and **B** $\in \mathcal{B}(\succ\succ)$. Alternatively, suppose **A** $\in \mathcal{B}(\succ\succ)$. Then **D** $\in \mathcal{G}(\succ\succ)$, **C** $\in \mathcal{B}(\succ\succ)$ and **B** $\in \mathcal{G}(\succ\succ)$. Both the partitions of {**A**, **B**, **C**, **D**} are admissible. ■

	L	*R*
T	1, 1, -5	-1, -5, 0
B	-5, -5, 0	0, 0, 10

ℓ

	L	*R*
T	-3, -1, 5	-5, -5, 0
B	2, -5, 20	-2, -2, 0

r

Figure 1

	L	R
T	1, 1, -5	-1, -5, 0
B	-5, -5, 0	0, 0, 10

ℓ

	L	R
T	-1, -1, 5	-5, -5, 0
B	-5, -5, 0	-2, -2, 0

c

	L	R
T	-3, -1, 5	-5, -5, 0
B	2, -5, 4	-2, -2, 0

r

Figure 2.

	L	C	R
T	1, 1, -5	-1, -5, 0	0, 0, 0
B	-5, -5, 0	0, 0, 10	0, 0, 0

ℓ

	L	C	R
T	-3, -1, 5	-5, -5, 0	0, 0, 0
B	2, -5, 20	-2, -2, 0	0, 0, 0

r

Figure 3.

	L	*R*
T	0, 5	0, 6
B	6, 0	1, 1

Figure 4.

	L	R
T	10, 10, 10	-3, 10.5, 4
B	9, 9, 6	11, 11, 5

ℓ

	L	R
T	9, 8, 3	-4, -5, 0
B	8, 12, 7	-2, -2, 11

r

Figure 5.

	L	R
T	10, 10, 10	-3, 9, 4
B	9, 9, 6	11, 11, 5

ℓ

	L	R
T	9, 8, 3	-4, -5, 0
B	8, 12, 7	-2, -2, 11

r

Figure 6.

U

ℓ

	L	*R*
T	10, 10, 10, 10	-3, 9, 4, -3
B	9, 9, 6, 9	11, 11, 5, 11

r

	L	*R*
T	9, 8, 3, 9	-4, -5, 0, -4
B	8, 12, 7, 8	-2, -2, 11, -2

D

ℓ

	L	*R*
T	12, 13, -1, 9	9, 13, -1, 9
B	9, 12, -1, 9	10, 13, -1, 9

r

	L	*R*
T	12, 10, -1, 9	9, 10, -1, 9
B	9, 10, 9, 9.5	9, 10, -1, 9

Figure 7.

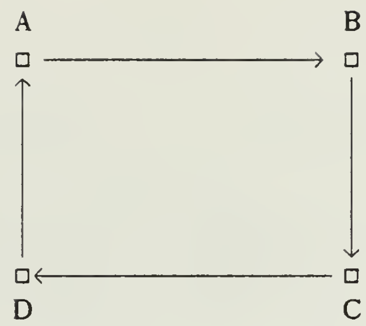

Figure 8.